DOCUMENT RESUME

ED 094 344 CS 001 242

AUTHOR Harris, Albert J.
TITLE Some New Developments in Readability.
PUB DATE Aug 74
NOTE 10p.; Paper presented at the Annual International Reading Association World Congress on Reading (5th, Vienna, Austria, Aug. 12-14, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Cloze Procedure; Linguistics; *Literature Reviews; *Readability; Reading; Reading Difficulty; *Reading Level; Reading Materials; *Reading Research; Vocabulary

ABSTRACT

One can measure the readability of any piece of reading material for an individual or a group by having the selection read and then testing for comprehension. Increasingly the cloze procedure has come to be used in preference to multiple-choice questions. Other readability scales have been based on previously scaled passages, on carefully graded books, and on the combined judgments of a group of experts. Two variables have consistently stood out as providing the best combination in the measurement of readability. The first is the difficulty of the vocabulary used, which is usually measured by finding the percentage of words that do not appear in a specific list of common, easy words. Spelling patterns as indices of vocabulary difficulty have just begun to be explored and seem promising. The second widely used variable is average sentence length, which seems to represent the many specific reasons why beginnings have been made in the automatic computer scoring of complicated linguistic variables such as syntactic depth and density. The specific factors that make some reading material hard to understand, such as vagueness, ambiguity, and lack of explicitness, are areas in which more research is needed. (WR)

Albert J. Harris
35 Rockwood Place
New Rochelle, N.Y. 10804, U.S.A.

SOME NEW DEVELOPMENTS ON READABILITY

(Presented at the Fifth IRA World Congress on Reading, Vienna, Austria, August 13, 1974)

One of the major tasks for the teacher of reading is to provide the student reader with material that he can read and understand. The term readability refers to the qualities of reading material which determine how easy it is to understand, how fluently it can be read, and how much interest it generates. This paper, however, is concerned only with readability in the sense of ease or difficulty of comprehension.

The term "new development" needs to be made more explicit. Two fairly recent publications provide a starting point. One is a revised annotated bibliography published by IRA in 1971 entitled Readability and Reading, compiled by Seels and Dale. This provides good coverage of research published between 1965 and 1970. The other is John Gilliland's Readability, a Teaching of Reading Monograph of the U.K.R.A., published in 1972. The present paper emphasizes developments too recent to have been covered in those two very helpful sources. Some earlier work will also be discussed.

On first thought it may seem that the logical way to determine the readability of a selection is to give it to someone to read and to check his comprehension. Such a procedure, sometimes called "trying the book on for size," can establish how well that person has been able to read it. But usually we want to predict how a selection, or a whole book, will fit an individual or group's ability to understand before deciding if they should be asked to read it.

The Measurement of Readability

To get away from the need to try a book with each potential user it is necessary to obtain a score which can express the book's readability on an easily interpreted scale. The way this is done is to measure its scores on characteristics which can predict where it would belong on a scale of reading selections whose readability scores have already been established.

The scale of reading selections that has been used most widely in readability research is a collection called Standard Test Lessons in Reading by McCall and Crabbs, originally published in the 1920's and most recently revised in 1961. There are five booklets, each containing about 70 one-page test lessons. For each lesson there is a short selection, followed by

about ten multiple-choice questions. At the bottom is a table giving the grade score corresponding to each possible number of right answers. These grade scores were originally established by giving the test lessons and a standardized reading test to several thousand pupils. The readability formulas developed by Lorge, Flesch, and Dale and Chall all used the McCall-Crabbs test lessons as a criterion.

Another way to set up a readability scale is to use the cloze technique to determine the comprehension difficulty of a number of selections. In the cloze procedure, words are deleted in a regular pattern and the reader's task is to write in each missing word. Usually every fifth word is deleted. The average number of correct cloze answers for a group can be used to rank selections in readability, setting up a scale against which other selections can be compared. The cloze technique has been used in much recent readability research, particularly by Bormuth (1968, 1969) and by Coleman (1968a, 1968b, 1971).

A third procedure is to take several series of carefully graded books and use their average characteristics to establish a graded scale. In the United States, widely used basal reader series have been used as criteria for the development of readability formulas by Spache and in the new Harris-Jacobson formulas which will be described a little later in this paper.

A fourth procedure is to have the selections rated by a group of judges, using the average rating to establish a scale. However, the ratings of judges on the same book can vary widely. Jongsma (1972) had 12 Newbery Award-winning books rated by 44 school and public librarians. One book was rated all the way from third grade to twelfth grade in difficulty.

Elements of Readability

Numerous studies have shown that the two main elements of the difficulty of reading material are the difficulty of the vocabulary used, or semantic difficulty, and the difficulty of the sentence structure, in other words, its syntactical or grammatical complexity. Of these two factors, vocabulary difficulty has consistently shown somewhat greater importance. Klare (1968) wrote: "Frequency of occurrence of words . . . clearly plays an all-pervasive role in language usage. Not only do humans tend to use some words much more often than others, they recognize more frequent words more rapidly than less frequent, prefer them, and understand and learn them more readily. It is not surprising, therefore, that this variable has such a central role in the measurement of readability."

Measurement of the difficulty of words has been done in a variety of ways. The procedure most often used is to determine what per cent of words in a sample are not found in a particular list of common words. Two lists compiled by Edgar Dale, a short list of 769 words and a list of about 3,000 words known to fourth grade children, have been used in several readability formulas including those by Lorge (1944), Dale and Chall (1948), and Spache (1953). Recently there have been efforts to

make these lists more up-to-date. Spache (1974) expanded the original short Dale list to 1,041 words, adding words from the Harris-Jacobson Core List (1972) and from library books of first and second grade levels, and revised his readability formula accordingly.

The publication in 1972 of the Harris-Jacobson Basic Elementary Reading Vocabularies provided a possible improvement in the measurement of vocabulary difficulty. The Harris-Jacobson word lists are based on a computerized analysis of about 4,500,000 running words found in 14 series of elementary school textbooks, from beginning first grade through sixth grade. A first grade list, a combined first and second grade list, and a combined first, second, and third grade list were tried out, and the combined first and second grade list proved to have the highest correlation with reading level from grade one through grade six. This combined list contains 912 root words and 1880 inflected forms, such as plurals and regular verb endings such as -ed and -ing, totalling 2,796 words. Since all allowable variants of a common root words are right in the list, one does not have to remember a variety of rules as to what constitutes an unfamiliar word, or repeatedly consult the list of rules. If a word is not in the list and is not a proper noun it counts as unfamiliar. The percent of unfamiliar words has the highest correlation with basal reader level of a large number of measures of vocabulary and sentence difficulty tried out.

Several other measures of vocabulary difficulty have been used in one or more of the more than fifty different readability formulas that have been published to date. Average number of letters per word, average number of syllables, average number of vowel letters, number of prefixes and suffixes, per cent of one-syllable words, per cent of words having three or more syllables, and per cent of words having more than five letters have been tried. Studies have also shown that the proportion of certain parts of speech is related to readability. Coleman (1971), for example, used number of pronouns and number of prepositions per 100 words as minor elements in his readability formulas.

All of these different ways of getting at vocabulary difficulty have substantial to high inter-correlations, but not all are of equal value in measuring readability. Unpublished results by Harris and Jacobson show that over the range from beginning first grade through sixth grade, per cent of unfamiliar words had the highest correlation, .87; per cent of words with more than five letters, .80; and average number of letters per word, .74. The per cent of unfamiliar words was 21 per cent more accurate in measuring readability than average number of letters per word. In recent years ways have developed for scoring many of these variables by computer. A computer can easily determine whether a word is or is not in a given list, count its letters, and so on, but it is much harder to write a program which will accurately identify syllables or prefixes. In the future only variables which can be scored by computer are likely to be retained in readability research.

In 1973 Harris and Jacobson reported that the per cent of words beginning with the letter e has a substantial correlation with difficulty in primary reading material, and Jacobson followed this up by identifying more than 1,000 spelling patterns in English words and correlating them with difficulty. In a 1974 paper Jacobson has reported that 37 spelling patterns, when combined in a multiple regression equation, correlated .92 with primary reading difficulty. In further work not yet published he has found that when one group of 12 spelling patterns is used for primary material and another group of 12 spelling patterns is used for middle-grade material, a combined multiple correlation of .965 was obtained. Using these spelling patterns is possible only with a computer and a very complicated computer program. Spelling patterns are probably related to readability in two different ways. One is that certain spelling patterns have greatly variable sound-symbol relationships. Initial e, for example, can represent at least seven different phonemes, as in each, ear, early, elephant, English, eight, and eyesight. This obviously increases the difficulty of decoding words beginning with e. The other way is that some spelling patterns tend to occur mainly in long words which are also uncommon and difficult, while other patterns appear predominantly in short, common, easy words. Spelling patterns present a new and promising approach to the measurement of vocabulary difficulty, but are usable only by those who have appropriate computer resources.

Sentence Difficulty

From the beginning of readability study is has been recognized that the difficulty of a sentence involves elements beyond the difficulties of the words in it. The average number of words per sentence has been used in many readability formulas and provides a reasonably satisfactory measure of those sentence characteristics that influence readability. In general, long sentences tend to have more modifiers and qualifiers, more embedded phrases and clauses, and complex rather than simple structure.

Since 1960 a number of efforts have been made to measure the linguistic or syntactic difficulty of sentences. Hunt (1965) developed a measure called the T-unit, which is one main clause and its related words, phrases, and clauses; in most cases, this amount to the number of words in the sentence. Golub (1969) developed a Syntactic Density Score which uses the T-unit and also takes account of such items as complex verb expansions and prepositional phrases. Since then Golub has developed a program which can obtain his Syntactic Density Score by computer(unpublished). Bormuth (1969) included measures of right depth of sentences and left depth of sentences among the variables used in some of his readability formulas. Botel and Granowsky (1972) developed a formula for measuring the syntactic density of individual sentences; the average score for the sentences in a selection should correlate with readability.

Such methods of sentence analysis obviously help to show why one sentence may be harder to understand than another of similar length. But there is no evidence as yet that they can give a better indication of the difficulty of a whole selection than average sentence length does.

MacGinitie and Tretiak (1972) compared the sentence depth measures devised by Yngve (1960) and by Allen (1968) with average sentence length and found that average sentence length gave the best prediction of readability when each was combined with a measure of word difficulty. Similar research is needed for the other measures of sentence complexity. Since scoring sentences for these variables is slow and laborious, future research in this area is likely to emphasize variables that can be scored by computer.

To sum up this discussion on the elements of comprehensibility or readability, the most important element in the difficulty of reading material is the difficulty of the vocabulary employed. This is measureable in a variety of ways, among which the per cent of words not found in a list of common, easy words seems to be the most satisfactory as well as the one most commonly used. The other main variable affecting readability is sentence complexity, which is well represented by average sentence length.

New Readability Formulas

The increasing use of computers in readability research has made it possible to develop readability formulas for special purposes with ease. However, most users of readability measures do not have computers available, and ease and speed of scoring and computation are important factors in the choice of which formula to use. This discussion will be confined to new formulas that are applicable by hand to children's materials for the elementary grades.

At the Denver Convention of IRA in May, 1973, Harris and Jacobson presented some data on new readability formulas. Since then the formulas have been further revised and a large number of new formulas have been tried out. Many different combinations of variables provide measures of about equal accuracy. Two formulas, one for grades one through three, and one which works from grade one through grade six but is mainly for use above third grade, have been selected as providing the best combination of ease and rapidity of application and high correlations with difficulty. Both formulas have correlations of .90 with reader level. The primary formula has a standard error of estimate of .38 of a year; the other formula has a standard error of .71 of a year. Complete directions for using these formulas will appear in a book scheduled for publication early in 1975 (Harris and Sipay, in press).

Spache (1974) published a revision of his formula for primary-grade material. The new formula is based on 100 samples of about 100 words each, or a total of about 10,000 words, in comparison to the Harris-Jacobson formulas which are based on 661 samples totalling about 135,000 words. Spache reports the very high correlation of .95 with book grade level.

Specific Features of Difficult Prose

The kinds of conjunctions used by writers seems to influence the difficulty of their material. Stoodt (1972) reported on the comparative difficulty of commonly used conjunctions for fourth grade children.

The ones that were best understood were and, for, and as. The following were comparatively difficult: when, so, but, or, where, while, that, and if. Robertson (1968) found that clauses introduced by however, thus, which, although, and yet were difficult for children in grades 4-6, and success in understanding such clauses was closely related to general reading comprehension.

Rosenshine (1969) used long passages which were equivalent in difficulty according for readability formulas but on which college students showed varying degrees of comprehension. He found five factors that tended to influence the readability of the passages. Difficulty was increased by vagueness and ambiguity, which is shown by the excessive use of indeterminate qualifiers such as rather, any number of, quite a bit, etc., and also by excessive use of probability words such as might, possibly, and sometimes. A factor aiding comprehension is frequent use of explaining links like because, in order to, if.., then, and so forth, which call attention to a cause, a result, or a means. A third factor was frequent use of examples, which seemed more important in difficult material than in easy material. The fourth factor was use of a rule-example-rule pattern, in which a generalization would be stated, followed by one or more examples, and then by a restatement of the generalization or rule. This seemed more beneficial than either an inductive or deductive pattern of presentation. The fifth factor was irrelevancy, which increased the difficulty of the material. Digressions and unnecessary restatements seem to lower the amount of information gain. These five factors are worth noting by those who write and lecture.

We have only a beginning of research on the specific features of prose writing that make material easy or difficult to read. We need to distinguish between inherent difficulty that results from the necessary use of concepts and relationships that are hard to explain and hard to understand, and unnecessary difficulty which is created by distinctive features of an author's style. When the subject-matter is inherently difficult, one can lower a readability formula score artificially by chopping long sentences into short ones and by substituting more common for less common synonyms. Geyer and Carey (1972) rewrote American History materials so as to reduce the Dale-Chall readability scores by about two grades. However, the comprehension scores of students were no higher on the easier than on the harder original versions. Apparently the kind of rewriting that was guided by the formula did not reduce the inherent difficulty of the content.

All too often, an author creates unnecessary difficulty for his readers. In addition to the factors of vagueness and irrelevancy discussed by Rosenshine, we have to guard against excessive use of the passive voice and the subjunctive mood. We have to check to see if modifying phrases and clauses are placed close to the items they modify. We have to note whether or not the antecedents of pronouns are easily identified. We have to observe whether sentences follow in a logical sequence, whether each paragraph actually has a main idea, whether appropriate emphasis is given to the most important statements. These are all factors that are not measured by readability formulas, but which do affect the ease or difficulty with which one can grasp the thought content of written material.

Readability of Languages Other Than English

This paper has considered only readability research on English prose. Readability studies have been conducted on quite a number of other languages, including French, German, Dutch, Spanish, Hebrew, Hindi, Chinese, Korean, Japanese, and Vietnamese. The limitations of this paper do not permit going into detail on these studies, most of which were done before 1970. Professor G. DeLandsheere of the University of Liege has been a leader in European research on readability. Very briefly, it seems safe to generalize that the two variables of vocabulary difficulty and sentence length are useful indicators of comprehensibility in any language. The features of specific written languages may encourage use of such variables as syllable counts or measures of the visual complexity of Chinese characters.

Summary

One can measure the readability of any piece of reading material for an individual or a group by having the selection read and testing for comprehension. Increasingly the cloze procedure has come to be used in preference to multiple-choice questions. A scale of selections can be arranged based on the average comprehension scores of a group. Other readability scales have been based on previously scaled passages, on carefully graded books, and on the combined judgments of a group of experts. Once a scale has been constructed, many different characteristics of the material can be measured, and each set of measurements can be correlated with the scale. Using the technique of multiple correlation, the best combination of variables and the best ways to weight them so as to get a maximum correlation can be discovered. Computer procedures which can score the selections for many variables as well as compute multiple correlations almost instantly have greatly simplified and speeded up readability research in recent years.

Consistently two variables have stood out as providing the best combination in the measurement of readability. The first, and more important, is the difficulty of the vocabulary used. This is usually measured by finding the per cent of words that do not appear in a specific list of common, easy words. Spelling patterns as indices of vocabulary difficulty have just begun to be explored and seem promising. The second widely used variable is average sentence length, which seems to represent the many specific reasons why long sentences are usually harder to understand than short ones. Now that beginnings have been made in the automatic computer scoring of complicated linguistic variables such as syntactic depth and density, more refined measures of sentence difficulty may appear in future readability formulas. Recent formulas that use the two variables of per cent of unfamiliar words and average number of words per sentence have multiple correlations of .90 or better with difficulty, seeming to indicate that not much further improvement is possible.

The specific factors that make some reading material hard to understand have begun to be analyzed. Vagueness, ambiguity, and lack of explicitness

concerning important relationships are among the factors noted thus far. Much more research is needed in this area.

Most readability research has been conducted on English prose. Some work has been done in a number of other languages, both European and non-European, which lies outside the scope of this paper.

Progress in the measurement and understanding of readability will improve the ability of authors and teachers to achieve a better fit between the abilities of readers to understand and the materials they are expected to read.

REFERENCES

Allen, Robert L. The structure of the English sentence. New York: Noble and Noble, 1968.

Bormuth, John R. (Ed.) Readability in 1968. A research bulletin prepared by a committee of the National Conference on Research in English. Urbana, Illinois: National Council of Teachers of English, 1968.

Bormuth, John R. Development of readability analyses. Washington, D.C.: Project No. 7-0052, Final Report, U.S. Office of Education, 1969.

Botel, Morton, and Granowsky, Alvin. A formula for measuring syntactic complexity: a directional effort. Elementary English, 1972, 49, 513-516.

Coleman, E.B. Experimental studies of readability, Part I: stimulus dimensions that affect readability. Elementary English, February 1968, 45, 166-178.(a)

Coleman, E.B. Experimental studies of readability, Part II: measures of readability and relevant populations. Elementary English, March 1968, 45, 316-323, 333. (b)

Coleman, E.B. Developing a technology of written instruction: some determinants of the complexity of prose. In E.Z. Rothkopf and P.E. Johnson (Eds.). Verbal learning and the technology of written instruction. New York: Teachers College Press, Columbia University, 1971.

Dale, Edgar, and Chall, Jeanne S. A formula for predicting readability. Educational Research Bulletin (Ohio State University), January 21 and February 17, 1948, 27, 11-20, 37-54.

De Landsheere, G., Henry, G., and Donnay, J. Étude sur la compréhension des messages contenus dans les journaux parlés et les journaux télévisés. Scientia Paedagogica Experimentalis, 1970, 7, 23-64.

Flesch, Rudolph. Marks of a readable style. New York: Teachers College Press, Columbia University, 1943.

Geyer, James R., and Carey, Albert R. Predicting and improving comprehensibility of social studies materials: the role of cloze procedure and readability adjustment. Reading World, December 1972, 12, 85-93.

Gilliland, John. Readability. London: University of London Press, 1972.

Golub, Lester S. Linguistic structures in students' oral and written discourse. Research in the Teaching of English, Spring 1969, 3, 70-85.

Harris, Albert J. and Jacobson, Milton D. Basic elementary reading vocabularies. New York: Macmillan Publishing Co., 1972.

Harris, Albert J., and Jacobson, Milton D. The Harris-Jacobson primary readability formulas. Paper presented at the Annual Convention of the International Reading Association, Denver, May 1973.

Harris, Albert J., and Sipay, Edward R. How to increase reading ability, Sixth edition revised. New York: David McKay Co., in press. Appendix D.

Hunt, Kellog. Grammatical structures written at three grade levels. NCTE Research Report No. 3. Urbana, Illinois: National Council of Teachers of English, 1965.

Jacobson, Milton D. Predicting reading difficulty from spelling. Spelling Progress Bulletin, Spring 1974, 14, 8-10.

Jongsma, Eugene A. The difficulty of children's books: Librarians' judgments vs. formula estimates. Elementary English, January 1972, 49, 20-26.

Klare, George R. The role of word frequency in readability. Elementary English, January 1968, 45, 12-22.

Lorge, Irving. Readability formulas -- an evaluation. Elementary English, February 1949, 26, 85-95.

MacGinitie, Walter H., and Tretiak, Richard. Sentence depth measures as predictors of reading difficulty. Reading Research Quarterly, Spring 1971, 6, 364-377.

McCall, William A. and Crabbs, Leila M. Standard test lessons in reading, revised. New York: Teachers College Press, Columbia University, 1961.

Robertson, Jean E. Pupil understanding of connectives in reading. Reading Research Quarterly, Spring 1968, 3, 387-417.

Lorge, Irving. Predicting readability. Teachers College Record, March 1944, 45, 404-419.

Rosenshine, Barak. New correlates of readability and listenability. In J. Allen Figurel (Ed.). Reading and Realism. Proceedings of the International Reading Association, 1969, 13, Part 1, 710-716.

Seels, Barbara, and Dale, Edgar, compilers. Readability and reading: an annotated bibliography, 1971 Revision. Newark, Delaware: International Reading Association, 1971.

Spache, George D. A new readability formula for primary-grade reading materials. Elementary School Journal, March 1953, 53, 410-413.

Spache, George D. Good reading for poor readers, Revised ninth edition, 1974. Champaign, Illinois: Garrard Publishing Co, 1974, 195-207.

Stoodt, Barbara D. The relationship between understanding grammatical conjunctions and reading comprehension. Elementary English, April 1972, 49, 502-504.

Yngve, V.H. A model and an hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 44, 444-466.